

Automated Learning and Data Analysis Final Report for Project

Due Apr 26th, 2019

**Dr. Thomas Price
Department of Computer Science**

Your final report is expected to be an 8-page report in the form of a NIPS paper. Your final project report should clearly articulate the rationale behind your project, in addition to simply reporting what you did and the results that were achieved. It should roughly have the following format:

- 1. Background**
- 2. Proposed method**
 1. Intuition - why should it be better than the state of the art?
 2. Description of its algorithms
- 3. Plan & Experiment**
 1. Dataset(s): Describe your dataset; number of features, number of rows, missing data, pre-processing procedure if any, and so on
 2. Description of your testbed; list of questions your experiments are designed to answer
 3. Details of the experiments
- 4. Results**
- 5. Conclusions**

The maximum page is 8 pages including all references. (Not including your Appendix). Upload your code to [GitHub](#) and include a link at the end of your final report.

The main objective for the final report paper is to let you have the opportunity to do independent research-oriented work by reading and analyzing research papers, surveying current frontiers of some research topics, and experimenting with his/her own new ideas on data mining. The final report paper will be evaluated based on the following factors:

1. The novelty of the paper. Final report papers reporting the student's own new ideas are favored in comparison to papers analyzing existing works.
2. The depth of the understanding and the critique on existing works. Papers involving good analysis and critique of existing works are favored in comparison to papers that only describe the existing works.
3. The writing style and the clarity of the written paper.

Below are more detailed things that we will be looking for when grading the project final reports. Note that this scheme is not intended to be a binding contract for exactly how I will mark the final project! It is only meant to give you guidance on the elements that I

will typically be looking for. The specific marking scheme will be adapted to the different types of projects that are submitted.

1. Background

- **Problem:** As clearly as possible, describe the general problem you are working on, and explain the specific refinements and special cases that you addressed.
- **Literature Survey:** Briefly survey the existing work that has been done on your problem, as well as the existing work that has been done on the approach (es) you are considering. You should cite some relevant references. Your survey need not be exhaustive, but you should try to cover the most important prior work if you can.

2. Methods

- **Approach:** As clearly as possible, describe the statistical technique(s) that you applied to the problem, and clearly specify the final system(s) that were implemented.
- **Rationale:** Explain why you implemented the systems that you did. Specifically, explain the simplifications you made along the way, and why you made these particular choices (and not others). (For example, you might detail other approaches that could be applied to the problem and explain why these were not pursued.)

3. Plan & Experiment

The most important (and most difficult) part of any research project is figuring out how to evaluate the results. You should state concrete goals for your project. In particular, you should identify at least one concrete hypothesis that you think would be interesting to test with your implemented system, and outline a specific plan for experiments that you would need to conduct in order to test each hypothesis. Testing it should be stated as the concrete goals for your project.

Note: You should not phrase your project goals as "my proposed approach will work great at solving problem X". Rather, you should phrase your goals as tests of specific hypotheses whose outcome would be interesting no matter what happened (so that your project would succeed however the tests turned out).

- **Datasets:** describe all the relevant information of the dataset that we need to know to understand your project.
- **Hypotheses:** Clearly state the main questions that you investigated in this project. These should be identified before you run the experiments! Ideally, these should be interesting questions whose answer is not obvious beforehand, but the answer would be interesting no matter how the experiments turned out.
- **Experimental Design:** Once you have settled on some good questions, it is important to figure out how to answer them. For each main question, describe a

series of experiments that is designed to answer the question. Explain the difficulties faced in designing these tests and explain why your experiments will overcome these difficulties to yield a definitive answer.

4. Results

- **Results:** As clearly as possible, describe the results you obtained with your tests. Use plots, graphs, and tables, if necessary. Make sure it is easy to understand what happened.
- **Critical Evaluation:** In the end it is important to critically evaluate the results of the experiments. Did the experimental results answer the questions? If so, then what were the answers? If not, then why did the results fail to yield definitive answers? Is it then possible to formulate a new experimental strategy?

Note that it is important to demonstrate critical thought in this part of your assessment. You will be graded more on the strength of your reasoning rather than how the tests actually turned out. For example, it is perfectly acceptable if you did not achieve definitive answers to your questions, so long as you can recognize this, explain why it happened, and suggest additional tests that might yield more definitive results.

5. Conclusions

- **Lessons Learned:** What specific things did you learn from doing this project? Did you learn anything about the problem itself, the approaches you tried, or the experiments you conducted? If there are any good ideas you came up with in the end but did not have time to pursue, this would be a good place to mention them.